



Chebyshev model arithmetic for factorable functions

Jai Rajyaguru¹ · Mario E. Villanueva¹ ·
Boris Houska² · Benoît Chachuat¹

Received: 13 March 2016 / Accepted: 4 October 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract This article presents an arithmetic for the computation of Chebyshev models for factorable functions and an analysis of their convergence properties. Similar to Taylor models, Chebyshev models consist of a pair of a multivariate polynomial approximating the factorable function and an interval remainder term bounding the actual gap with this polynomial approximant. Propagation rules and local convergence bounds are established for the addition, multiplication and composition operations with Chebyshev models. The global convergence of this arithmetic as the polynomial expansion order increases is also discussed. A generic implementation of Chebyshev model arithmetic is available in the library MC++. It is shown through several numerical case studies that Chebyshev models provide tighter bounds than their Taylor model counterparts, but this comes at the price of extra computational burden.

Keywords Global optimization · Factorable functions · Chebyshev models · Taylor models · Interval analysis · Convergence rate

1 Introduction

Many complete search methods for problems in global optimization and constraint satisfaction hinge on the ability to compute enclosures for the range of nonconvex functions as well as sets defined by multiple equalities and/or inequalities [44]. A variety of parameterizations can be used to describe or approximate compact sets, including convex sets such as intervals, ellipsoids and polytopes as well as nonconvex sets for instance in the form of the image of a multivariate polynomial. Chachuat et al. [13] have argued that the foregoing parameter-

✉ Benoît Chachuat
b.chachuat@imperial.ac.uk

¹ Department of Chemical Engineering, Centre for Process Systems Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

² School of Information Science and Technology, ShanghaiTech University, 319 Yueyang Road, Shanghai 200031, China

izations and many others can be cast as affine set-parameterizations with either convex or nonconvex basis sets. In particular, a number of arithmetics have been developed over the years that enable propagation of certain affine-set parameterizations through the atom operations of factorable functions in order to determine rigorous enclosures of their image sets. This includes interval arithmetic [40,41], ellipsoidal calculus [28,61], polyhedral relaxations [54,55], and Taylor model arithmetic [11,36], to name but a few. Other established techniques to relax sets or programs defined by nonconvex factorable functions involve constructing a pair of convex/concave bounds for each function, such as McCormick relaxations [38,39] and α BB/ γ BB relaxations [1,3,23]. In the case of semialgebraic problems, hierarchies of linear matrix inequality (LMI) relaxations can also be constructed based on a semidefinite programming formulation of the sum of squares decomposition for multivariate polynomials [29,46].

The focus in this paper is on *polynomial models* of multivariable functions, namely enclosures in the form of a polynomial approximant of given order and an interval remainder bound on the approximation error. Taylor models (or Taylor forms) are a special case of polynomial models, whereby the polynomial approximation matches the multivariate Taylor expansion of the function at a given point in the variable domain [43]. The idea of such interval polynomial enclosures dates back to Moore [40], and methods for computing the remainder bounds concurrently with the polynomial were later developed in the early 1980s [17,51] and popularized from the mid 1990s by Berz et al. [7,35,36]. Similar in essence to interval arithmetic, a well-developed Taylor model arithmetic is available [8,34], encompassing rules for binary sums, binary products and outer-compositions with a library of univariate functions such as $\exp(\cdot)$, $\log(\cdot)$ and $\sqrt{\cdot}$. For sufficiently smooth functions, the diameter of the remainder interval constructed with (q th-order) Taylor model arithmetic can be proved to be a high-order power ($q + 1$) of the diameter of the variables domain [11]. This gives Taylor models a clear advantage over traditional interval extensions or centered forms for sufficiently narrow domains, but conversely it may result in a large overestimation or may even be poorer than naive interval evaluation over wider domains. Nevertheless, this approach has proved successful in computing tight enclosures for the solutions of differential equations and implicit algebraic equations [24,33,42,49,50,53,60], and it has enabled complete search for a range of global optimization or constraint satisfaction problems that could not be tackled using interval techniques alone (see, e.g., [4,9,31,32,47,52]). Such higher-order inclusion techniques are indeed appealing in complete search applications based on branching or subdivision, where they can mitigate the clustering effect [15,62].

A rather natural idea involves replacing the Taylor series approximation with a better one, ideally the minimax polynomial approximation, or near-minimax approximations such as Chebyshev series. The development of affine arithmetics [19] yielding first-order approximations can be seen as a precursor of this approach. Regarding univariate functions, improved variants based on the Chebyshev polynomials, Bernstein polynomials and others were widely used during the early 1980s under the name of ultra-arithmetic [26]. A comparison of univariate Taylor forms and Chebyshev forms in [27] showed that expansions in Chebyshev series may be orders of magnitude more accurate than expansions in Taylor series. The use of Chebyshev series expansion as an arithmetic is also the philosophy behind the package **Chebfun** by Trefethen et al. [6,57], which has recently been extended to functions in two variables as well [56]. It should be noted that **Chebfun** relies on computations with functions to 15-digit accuracy, thus removing the need for propagating a remainder term, but the results are not validated per se. In contrast, Brisebarre and Joldeş [12,25] developed an arithmetic for univariate functions based on Chebyshev interpolation polynomials and a remainder term. More recently, Dzetkulič [16] extended the previous work to enable encl-

sure of multivariate functions using truncated Chebyshev series and an interval remainder term, namely Chebyshev models. This latter work makes use of Chebyshev models as a means of propagating bounds for differential equations, but does not study the convergence properties of this arithmetic.

Building upon these existing contributions, this article presents a Chebyshev model arithmetic in order to compute enclosures for an inclusive class of factorable functions (Sect. 3). Besides formalizing the rules of addition, multiplication and composition as well as the polynomial range bounds, we describe an improvement of the remainder term in composition operations with certain univariate functions. We also present a convergence analysis of Chebyshev models, which investigates the propagation of local convergence properties through the addition, multiplication and composition operations (Sect. 4), and we discuss the global convergence of this arithmetic as the polynomial expansion order increases. Finally, we present numerical case studies illustrating the convergence properties of Chebyshev models and comparing their tightness and computational performance with Taylor models (Sect. 5). The underlying implementation of Chebyshev model arithmetic is part of the library MC++, which is made freely available at: <http://omega-icl.bitbucket.org/mcpp/>.

2 Notation and background

The set of n -dimensional interval vectors is denoted by \mathbb{IR}^n . The midpoint and radius of an interval vector $X := [x^L, x^U] \in \mathbb{IR}^n$ are defined as $\text{mid } X := \frac{1}{2}(x^L + x^U) \in \mathbb{R}^n$ and $\text{rad } X := \frac{1}{2}(x^U - x^L) \in \mathbb{R}^n$, respectively. The image of the set $X \in \mathbb{IR}^n$ under a continuous function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is given by

$$\bar{f}(X) := \{f(x) | x \in X\} \in \mathbb{IR}.$$

The total variation of a (possibly discontinuous) function $g: \mathbb{R} \rightarrow \mathbb{R}$ on the interval $Y := [y^L, y^U] \in \mathbb{IR}$ is the quantity

$$V[g]_Y := \sup \left\{ \sum_{i=1}^N |g(\xi_i) - g(\xi_{i-1})| \mid N > 0, \quad y^L = \xi_0 < \xi_1 < \dots < \xi_N = y^U \right\},$$

where ξ_0, \dots, ξ_N is any finite partition of Y , and the supremum norm of g on Y is given by

$$\|g\|_Y := \sup \{ |g(\xi)| \mid \xi \in Y \}$$

Moreover, by a slight abuse of the notation, we use $V[g]$ and $\|g\|$ when $Y = [-1, 1]$.

Chebyshev polynomials of the first kind, denoted by T_k , are defined by the three-term recurrence relation

$$T_0(x) := 1, \quad T_1(x) := x, \quad T_{k+1}(x) := 2xT_k(x) - T_{k-1}(x).$$

An explicit expression of these polynomials on $[-1, 1]$ is

$$T_k(x) = \cos(k \arccos(x)).$$

The Chebyshev polynomials are extremal polynomials with respect to the property that, for each $k > 0$, $T_k(x)$ is the polynomial with the largest possible leading coefficient subject to the condition that its range is in $[-1, 1]$ for $x \in [-1, 1]$.

Any polynomial function \mathcal{P}^q of degree q can be written in terms of the Chebyshev polynomials as [37]

$$\mathcal{P}^q(x) = \sum_{k=0}^q a_k T_k(x).$$

An efficient way of evaluating such a linear combination of Chebyshev polynomials is using the Clenshaw algorithm [14],

$$\mathcal{P}^q(x) = a_0 + xb_1(x) - b_2(x),$$

where the b_k 's are obtained using the (reverse) recurrence formula

$$b_k(x) = a_k + 2xb_{k+1}(x) - b_{k+2}(x) \quad \text{with} \quad b_{q+1}(x) = b_{q+2}(x) = 0. \quad (1)$$

Moreover, the product between two Chebyshev polynomials expands to

$$T_j(x)T_k(x) = \frac{1}{2} (T_{j+k}(x) + T_{|j-k|}(x)). \quad (2)$$

The Chebyshev polynomials form a sequence of orthogonal polynomials with respect to the weight $\frac{1}{\sqrt{1-x^2}}$ on $[-1, 1]$,

$$\int_{-1}^1 \frac{T_j(x)T_k(x)}{\sqrt{1-x^2}} dx = \begin{cases} \pi & \text{if } j = k = 0 \\ \pi/2 & \text{if } j = k > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Any Lipschitz continuous function f on $[-1, 1]$ has a unique representation as an absolutely and uniformly convergent series [58],

$$f(x) = \sum_{k=0}' a_k T_k(x) \quad \text{with} \\ a_k := \frac{2}{\pi} \int_{-1}^1 \frac{T_k(x)f(x)}{\sqrt{1-x^2}} dx = \frac{2}{\pi} \int_0^\pi f(\cos(\theta)) \cos(k\theta) d\theta, \quad (3)$$

where the notation \sum' indicates that the first term is halved. In general, the Chebyshev coefficients a_k can be approximated using a numerical quadrature, possibly in combination with the fast Fourier cosine transform to improve the computation speed [21].

For practical purposes, a q th-order polynomial approximant \mathcal{P}_f^q for f can be obtained from the partial sums of (3) as

$$\mathcal{P}_f^q(x) := \sum_{k=0}'^q a_k T_k(x). \quad (4)$$

The smoother the function f on $[-1, 1]$, the faster the convergence of its approximants \mathcal{P}_f^q as $q \rightarrow \infty$. The following bounds are available with regards to the accuracy of the partial sum \mathcal{P}_f^q , depending on the regularity of f :

- If f and its derivatives through $f^{(s-1)}$ for a given $s \geq 0$ are absolutely continuous on $[-1, 1]$ and if $f^{(s)}$ has finite total variation, $V[f^{(s)}] < \infty$, then [58]

$$\forall q \geq s+1, \quad \|f - \mathcal{P}_f^q\| \leq \frac{2}{\pi} \frac{V[f^{(s)}]}{s(q-s)^s}. \quad (5)$$

- If f and its derivatives through $f^{(q)}$ are continuously differentiable on $[-1, 1]$, then [18]

$$\|f - \mathcal{P}_f^q\| \leq \frac{\|f^{(q+1)}\|}{2^q(q+1)!}. \quad (6)$$

An alternative approach to obtaining a q th-order polynomial approximant is by constructing the polynomial interpolant $\widehat{\mathcal{P}}_f^q$ at Chebyshev points,

$$\begin{aligned} \widehat{\mathcal{P}}_f^q(x) &:= \sum_{k=0}^q \widehat{a}_k T_k(x) \\ \text{with } \widehat{a}_k &:= \frac{2}{q+1} \sum_{j=0}^q f(\cos(\theta_j)) T_k(\cos(\theta_j)) = \frac{2}{q+1} \sum_{j=0}^q f(\cos(\theta_j)) \cos(k\theta_j) \\ \text{and } \theta_j &:= \frac{\pi}{2} \frac{j+1}{q+1}. \end{aligned} \quad (7)$$

The polynomial $\widehat{\mathcal{P}}_f^q$ yields a near-minimax approximation, and counterparts to the error estimates in (5) and (6) are, respectively, [18, 58]

$$\forall q \geq s+1, \quad \|f - \widehat{\mathcal{P}}_f^q\| \leq \frac{4}{\pi} \frac{V[f^{(s)}]}{s(q-s)^s}, \quad (8)$$

$$\|f - \widehat{\mathcal{P}}_f^q\| \leq \frac{\|f^{(q+1)}\|}{2^q(q+1)!}. \quad (9)$$

While the truncated Chebyshev series expansion \mathcal{P}_f^q is generally a better approximation than the Chebyshev interpolation polynomial $\widehat{\mathcal{P}}_f^q$, the latter comes with the inherent advantage that the coefficients \widehat{a}_k in (7) can be obtained from a few function evaluations only.

The (over-)approximation of multivariate functions using Chebyshev polynomials is at the heart of the present paper. Like in the univariate case, any q th-order polynomial function in n variables can be written in terms of the Chebyshev polynomials as

$$\mathcal{P}^q(x) = \sum_{|\kappa| \leq q} a_\kappa T_\kappa(x),$$

where $\kappa \in \mathbb{N}^n$ is a multi-index; the order of κ is given by $|\kappa| := \sum_{i=1}^n \kappa_i$; and $T_\kappa(x)$ is a shorthand notation for the expression $\prod_{i=1}^n T_{\kappa_i}(x_i)$ for any $x \in \mathbb{R}^n$.

3 Chebyshev model arithmetic

Similar to Taylor models [11, 43], Chebyshev models are estimators made by the sum of a multivariate polynomial in Chebyshev form and a remainder interval providing a bound on the approximation error. A formal definition follows:

Definition 1 Let the function $f : Z \rightarrow \mathbb{R}$ be defined on $Z \subseteq \mathbb{R}^n$. For every interval vector $Y \subset Z$, let the multivariate polynomial $\mathcal{P}_{f,Y}^q : [-1, 1]^n \rightarrow \mathbb{R}$ with $\mathcal{P}_{f,Y}^q(\xi) := \sum_{|\kappa| \leq q} a_\kappa T_\kappa(\xi)$ and the symmetric interval $\mathcal{R}_{f,Y}^q \in \mathbb{IR}$ be such that

$$\forall \xi \in [-1, 1]^n, \quad f(\text{mid } Y + \xi \circ \text{rad } Y) - \mathcal{P}_{f,Y}^q(\xi) \in \mathcal{R}_{f,Y}^q, \quad (10)$$

where \circ denotes the entrywise (or Hadamard) product. We call the pair $(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q)$ a *qth-order Chebyshev model* of f on Y , and $(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q)_{Y \subseteq Z}$ a *scheme of qth-order Chebyshev models* for f in Z .

Notice that, in the sense of Definition 1, the polynomial part $\mathcal{P}_{f,Y}^q$ of a Chebyshev model may be different from the actual truncated Chebyshev expansion of f . Notice also the need for scaling the variables in $[-1, 1]$ for consistency with the definition and properties of Chebyshev polynomials of the first kind. The definition of schemes of Chebyshev models is in analogy with the scheme of estimators introduced by Bompadre et al. [10, 11] to support convergence analysis.

Similar in essence to interval (or Taylor model) arithmetic, our focus in the following subsections is on the development of an arithmetic that enables the construction of Chebyshev models for factorable functions, namely functions which can be represented by a finite number of binary sums, binary products and outer compositions with a univariate function. As well as the specification of these three “elementary” operations, developing such an arithmetic also calls for an initialization procedure for variables and constants, and the ability to bound the range of Chebyshev models.

3.1 Initialization

Given a set $Z \subseteq \mathbb{R}^n$, each variable $z_i \in Z_i$ with $i \in \{1, \dots, n\}$ can be represented by a scheme of *qth-order Chebyshev models* $(\mathcal{P}_{z_i,Y}^q, \mathcal{R}_{z_i,Y}^q)_{Y \subseteq Z}$ with

$$\forall \xi \in [-1, 1]^n, \quad \mathcal{P}_{z_i,Y}^q(\xi) := \begin{cases} \text{mid } Y_i \, T_0(\xi_i) & \text{if } q = 0 \\ \text{mid } Y_i \, T_0(\xi_i) + \text{rad } Y_i \, T_1(\xi_i) & \text{if } q > 0 \end{cases} \quad (11)$$

$$\mathcal{R}_{z_i,Y}^q := \begin{cases} Y_i - \text{mid } Y_i & \text{if } q = 0 \\ [0, 0] & \text{if } q > 0. \end{cases} \quad (12)$$

Likewise, any real constant $r \in \mathbb{R}$ is trivially represented by the scheme of Chebyshev models $(r \, T_0, [0, 0])_{Y \subseteq Z}$.

3.2 Binary sum operations

Consider two functions $f_1, f_2 : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, and suppose that $(\mathcal{P}_{f_1,Y}^q, \mathcal{R}_{f_1,Y}^q)_{Y \subseteq Z}$ and $(\mathcal{P}_{f_2,Y}^q, \mathcal{R}_{f_2,Y}^q)_{Y \subseteq Z}$ are corresponding schemes of Chebyshev models. Then, a scheme of Chebyshev models $(\mathcal{P}_{f_1 \pm f_2,Y}^q, \mathcal{R}_{f_1 \pm f_2,Y}^q)_{Y \subseteq Z}$ for $f_1 \pm f_2$ in Z is simply obtained by addition/subtraction of the polynomial parts and of the remainder parts,

$$\forall \xi \in [-1, 1]^n, \quad \mathcal{P}_{f_1 \pm f_2,Y}^q(\xi) := \mathcal{P}_{f_1,Y}^q(\xi) \pm \mathcal{P}_{f_2,Y}^q(\xi) \quad (13)$$

$$\mathcal{R}_{f_1 \pm f_2,Y}^q := \mathcal{R}_{f_1,Y}^q \pm \mathcal{R}_{f_2,Y}^q. \quad (14)$$

3.3 Binary product operations

Consider again two functions $f_1, f_2 : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, and suppose that $(\mathcal{P}_{f_1,Y}^q, \mathcal{R}_{f_1,Y}^q)_{Y \subseteq Z}$ and $(\mathcal{P}_{f_2,Y}^q, \mathcal{R}_{f_2,Y}^q)_{Y \subseteq Z}$ are corresponding schemes of Chebyshev models, with

$$\mathcal{P}_{f_1,Y}^q(\xi) := \sum_{|\kappa| \leq q} a_{Y,\kappa}^q T_\kappa(\xi) \quad \text{and} \quad \mathcal{P}_{f_2,Y}^q(\xi) := \sum_{|\kappa| \leq q} b_{Y,\kappa}^q T_\kappa(\xi).$$

In order to arrive at a valid scheme of q th-order Chebyshev models $(\mathcal{P}_{f_1 f_2, Y}^q, \mathcal{R}_{f_1 f_2, Y}^q)_{Y \subset Z}$ for the product function $f_1 f_2$ in Z , we start by expressing the product between the respective polynomial parts,

$$\mathcal{P}_{f_1, Y}^q(\xi) \mathcal{P}_{f_2, Y}^q(\xi) = \sum_{|\lambda| \leq q} \sum_{|\mu| \leq q} a_{Y, \lambda}^q b_{Y, \mu}^q T_\lambda(\xi) T_\mu(\xi).$$

A complication with polynomial multiplication in Chebyshev basis is that, according to the property (2), each product term $T_\lambda(\xi) T_\mu(\xi)$ generates $2^{N(\lambda, \mu)}$ terms with

$$N(\lambda, \mu) := \text{card}\{i \in \{1, \dots, n\} \mid \lambda_i \mu_i > 0\} \leq n.$$

See also Sect. 5 for further discussions and related implementation considerations.

Introducing the map $\mathbb{P}_q: \mathbb{N}^n \rightarrow \mathbb{N}^n \times \mathbb{N}^n$ given by

$$\mathbb{P}_q(\kappa) := \left\{ (\lambda, \mu) \in \mathbb{N}^{2n} \mid \forall i \in \{1, \dots, n\}, \quad \lambda_i + \mu_i = \kappa_i \vee |\lambda_i - \mu_i| = \kappa_i \right\}, \quad (15)$$

it follows from (2) that

$$\mathcal{P}_{f_1, Y}^q(\xi) \mathcal{P}_{f_2, Y}^q(\xi) = \sum_{|\kappa| \leq 2q} c_{Y, \kappa}^q T_\kappa(\xi) \quad \text{with} \quad c_{Y, \kappa}^q := \sum_{(\lambda, \mu) \in \mathbb{P}_q(\kappa)} \frac{a_{Y, \lambda}^q b_{Y, \mu}^q}{2^{N(\lambda, \mu)}}. \quad (16)$$

Observe that the resulting product polynomial is of order $2q$ and that the index sets $\mathbb{P}_q(\kappa)$ are non-empty for $|\kappa| \leq 2q$. A simple way of defining the polynomial part $\mathcal{P}_{f_1 f_2, Y}^q$ in the resulting Chebyshev model is therefore by retaining all the terms of order no larger than q as

$$\forall \xi \in [-1, 1]^n, \quad \mathcal{P}_{f_1 f_2, Y}^q(\xi) := \sum_{|\kappa| \leq q} c_{Y, \kappa}^q T_\kappa(\xi). \quad (17)$$

Then, the corresponding remainder term $\mathcal{R}_{f_1 f_2, Y}^q$ can be obtained by assembling four terms, namely the multiplication of the two remainders, the multiplication of each remainder with a bound on the other polynomial range, and a bound on all the terms of order larger than q from the product of the two polynomials,

$$\begin{aligned} \mathcal{R}_{f_1 f_2, Y}^q &:= \mathcal{R}_{f_1, Y}^q \mathcal{R}_{f_2, Y}^q + \mathcal{R}_{f_1, Y}^q \left[\mathcal{P}_{f_2, Y}^q \right] + \left[\mathcal{P}_{f_1, Y}^q \right] \mathcal{R}_{f_2, Y}^q \\ &\quad + \sum_{q < |\kappa| \leq 2q} \left| c_{Y, \kappa}^q \right| [-1, 1]. \end{aligned} \quad (18)$$

The latter hinges on the ability to compute bounds on the polynomial parts of the Chebyshev model operands, denoted by $[\mathcal{P}_{f_1, Y}^q] \supseteq \{\mathcal{P}_{f_1, Y}^q(\xi) \mid \xi \in [-1, 1]^n\}$ and $[\mathcal{P}_{f_2, Y}^q] \supseteq \{\mathcal{P}_{f_2, Y}^q(\xi) \mid \xi \in [-1, 1]^n\}$ here. Such range bounds for Chebyshev models are discussed later on in Sect. 3.5.

An important aspect about polynomial multiplication in Chebyshev basis is that the computed q th-order polynomial $\mathcal{P}_{f_1 f_2, Y}^q$ in (17) will not correspond to the q th-order Chebyshev expansion of $f_1 f_2$ in general, even when $\mathcal{P}_{f_1, Y}^q$ and $\mathcal{P}_{f_2, Y}^q$ are themselves q th-order Chebyshev expansion of f_1 and f_2 , respectively. The reason behind this discrepancy is that the Chebyshev coefficients in the q th-order expansion of $f_1 f_2$ depend on terms of order greater than q in the Chebyshev expansions of f_1 and f_2 , which are missing here because of the truncation. Dzetkulić [16] argue that the resulting Chebyshev models nevertheless provide a tighter approximation than by multiplying $\mathcal{P}_{f_1, Y}^q$ and $\mathcal{P}_{f_2, Y}^q$ in monomial basis.

3.4 Univariate outer-composition operations

Consider the functions $f : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $F : X \subseteq \mathbb{R} \rightarrow \mathbb{R}$. Suppose that $(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q)_{Y \subset Z}$ is a scheme of q th-order Chebyshev models for f , such that $B_Y^q := [\mathcal{P}_{f,Y}^q] + \mathcal{R}_{f,Y}^q \subseteq X$ for all $Y \subseteq Z$. Let the scaling function $\Phi_Y^q : Y \rightarrow [-1, 1]$ be given by

$$\Phi_Y^q(y) := \frac{f(y) - \text{mid } B_Y^q}{\text{rad } B_Y^q} \quad (:= 0 \quad \text{if } \text{rad } B_Y^q = 0),$$

and denote by $(\mathcal{P}_{\Phi,Y}^q, \mathcal{R}_{\Phi,Y}^q)_{Y \subset Z}$ the corresponding scheme of Chebyshev models for Φ_Y^q ,

$$(\mathcal{P}_{\Phi,Y}^q, \mathcal{R}_{\Phi,Y}^q) = \frac{(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q) - \text{mid } B_Y^q}{\text{rad } B_Y^q} \quad (= (0, [0, 0]) \quad \text{if } \text{rad } B_Y^q = 0), \quad (19)$$

with $\mathcal{P}_{\Phi,Y}^q(0) = 0$. Suppose also that a scheme $(\mathcal{P}_{F,W}^q, \mathcal{R}_{F,W}^q)_{W \subset X}$ of Chebyshev models for F is available, with

$$\forall \zeta \in [-1, 1], \quad \mathcal{P}_{F,W}^q(\zeta) := \sum_{k=0}^q \varphi_{W,k}^q T_k(\zeta).$$

We will discuss the computation of such a scheme at the end of this subsection.

In order to construct a valid scheme of q th-order Chebyshev models $(\mathcal{P}_{F \circ f,Y}^q, \mathcal{R}_{F \circ f,Y}^q)_{Y \subset Z}$ for the composite function $F \circ f$ in Z , we start by applying the Clenshaw recurrence formula (1) in order to determine a scheme $(\mathcal{P}_{\mathcal{P}_{F \circ \Phi},Y}^q, \mathcal{R}_{\mathcal{P}_{F \circ \Phi},Y}^q)_{Y \subset Z}$ of Chebyshev models for the (parameterized) composite function $\mathcal{P}_{F,B_Y^q}^q \circ \Phi_Y^q$,

$$(\mathcal{P}_{\mathcal{P}_{F \circ \Phi},Y}^q, \mathcal{R}_{\mathcal{P}_{F \circ \Phi},Y}^q) := \varphi_{B_Y^q,0}^q + \left(\mathcal{P}_{\Phi,Y}^q, \mathcal{R}_{\Phi,Y}^q \right) \left(\mathcal{P}_{\beta_1,Y}^q, \mathcal{R}_{\beta_1,Y}^q \right) - \left(\mathcal{P}_{\beta_2,Y}^q, \mathcal{R}_{\beta_2,Y}^q \right) \quad (20)$$

$$\text{with:} \quad \left(\mathcal{P}_{\beta_k,Y}^q, \mathcal{R}_{\beta_k,Y}^q \right) = \varphi_{B_Y^q,k}^q + 2 \left(\mathcal{P}_{\Phi,Y}^q, \mathcal{R}_{\Phi,Y}^q \right) \left(\mathcal{P}_{\beta_{k+1},Y}^q, \mathcal{R}_{\beta_{k+1},Y}^q \right) - \left(\mathcal{P}_{\beta_{k+2},Y}^q, \mathcal{R}_{\beta_{k+2},Y}^q \right), \quad k = 0, \dots, q \quad (21)$$

$$\left(\mathcal{P}_{\beta_{q+1},Y}^q, \mathcal{R}_{\beta_{q+1},Y}^q \right) = \left(\mathcal{P}_{\beta_{q+2},Y}^q, \mathcal{R}_{\beta_{q+2},Y}^q \right) = (0, [0, 0]). \quad (22)$$

Observe that this recurrence consists of binary sum and product operations only, for which rules have already been established in Sects. 3.2 and 3.3. Observe also that when $\text{rad } B_Y^q = 0$, which occurs if either f is constant on Y or $\text{rad } Y = 0$, we have

$$\left(\mathcal{P}_{\mathcal{P}_{F \circ \Phi},Y}^q, \mathcal{R}_{\mathcal{P}_{F \circ \Phi},Y}^q \right) := \left(\varphi_{B_Y^q,0}^q, [0, 0] \right).$$

Finally, the polynomial and remainder parts of the desired Chebyshev models $(\mathcal{P}_{F \circ f,Y}^q, \mathcal{R}_{F \circ f,Y}^q)_{Y \subset Z}$ are given by

$$\forall \xi \in [-1, 1]^n, \quad \mathcal{P}_{F \circ f,Y}^q(\xi) := \mathcal{P}_{\mathcal{P}_{F \circ \Phi},Y}^q(\xi) \quad (23)$$

$$\mathcal{R}_{F \circ f,Y}^q := \mathcal{R}_{\mathcal{P}_{F \circ \Phi},Y}^q + \mathcal{R}_{F,B_Y^q}^q. \quad (24)$$

In order to use this approach, one needs the ability to construct Chebyshev models $(\mathcal{P}_{F,W}^q, \mathcal{R}_{F,W}^q)$ corresponding to a given library of univariate functions, such as $F(x) = \exp(x)$, $\log(x)$, \sqrt{x} , $\frac{1}{x}$, $\sin(x)$, $|x|$, etc.

- A systematic way of doing this is by computing a truncated Chebyshev expansion of F of order q as

$$\mathcal{P}_{F,W}^q(\xi) = \sum_{k=0}^q \frac{2}{\pi} \left(\int_{-1}^1 \frac{T_k(\zeta) F(\text{mid } W + \zeta \text{ rad } W)}{\sqrt{1 - \zeta^2}} d\zeta \right) T_k(\xi). \quad (25)$$

Then, if F is $q + 1$ -times continuously differentiable on W , the truncation error can be bounded as

$$\mathcal{R}_{F,W}^q = \frac{\|F^{(q+1)}\|_W (\text{rad } W)^{q+1}}{2^q (q+1)!} [-1, 1]. \quad (26)$$

Otherwise, if F and its derivatives through $F^{(s-1)}$ are absolutely continuous on W and $F^{(s)}$ has finite total variation $V[F^{(s)}]_W < \infty$ with $s < q$, a bound on the truncation error is given by

$$\mathcal{R}_{F,W}^q = \frac{2}{\pi} \frac{V[F^{(s)}]_W (\text{rad } W)^s}{s(q-s)^s} [-1, 1]. \quad (27)$$

- Alternatively, a Chebyshev interpolating polynomial of order q may be computed as

$$\mathcal{P}_{F,W}^q(\xi) = \sum_{k=0}^q \frac{2}{q+1} \left(\sum_{j=0}^q F(\text{mid } W + \zeta_j \text{ rad } W) T_k(\zeta_j) \right) T_k(\xi). \quad (28)$$

The error bound $\mathcal{R}_{F,W}^q$ in (26) remains valid for any $q + 1$ -times continuously differentiable function F on W . Otherwise, the following bound may be used instead of (27) if F and its derivatives through $F^{(s-1)}$ are absolutely continuous on W and $F^{(s)}$ has finite total variation $V[F^{(s)}]_W < \infty$ with $s < q$,

$$\mathcal{R}_{F,W}^q = \frac{4}{\pi} \frac{V[F^{(s)}]_W (\text{rad } W)^s}{s(q-s)^s} [-1, 1]. \quad (29)$$

In comparison with their Taylor expansion counterparts for sufficiently smooth functions, we note that the width of the remainder is smaller by a factor of 2^q . Moreover, the Chebyshev approach remains applicable for continuous, yet nonsmooth, functions. For instance, since the total variation of the derivative of $|x|$ on the interval $[-\rho, \rho]$ is finite, equal to 2, a bound for the interpolation error based on (29) is $\mathcal{R}_{|\cdot|, [-\rho, \rho]}^q = [-\frac{8\rho}{\pi(q-1)}, \frac{8\rho}{\pi(q-1)}]$, for any $q > 1$.

In practice, the bounds $\mathcal{R}_{F,W}^q$ on the truncation error may be tightened further if F meets certain regularity and monotonicity conditions. Under these extra conditions, it can be established that the maximum absolute error between F and its polynomial approximant $\mathcal{P}_{F,W}^q$ —as given by (25) or (28)—always occurs at one of the endpoints of W . This result is formalized in the following lemma:

Lemma 1 *Let the function F have an analytic extension on the closed unit disk in the complex plane and have all its successive derivatives $F^{(k)}$, $k \geq 1$ either of the same or alternating sign in $[-1, 1]$. Then, we have*

$$\|F - \mathcal{P}_{F,W}^q\| = \max \{ |F(-1) - \mathcal{P}_{F,W}^q(-1)|, |F(1) - \mathcal{P}_{F,W}^q(1)| \},$$

$$\text{and } \|F - \widehat{\mathcal{P}}_F^q\| = \max \{ |F(-1) - \widehat{\mathcal{P}}_F^q(-1)|, |F(1) - \widehat{\mathcal{P}}_F^q(1)| \}$$

for any $q \geq 0$, with \mathcal{P}_F^q and $\widehat{\mathcal{P}}_F^q$ given by (4) and (7), respectively.

Proof See “Appendix 1”. □

This method of calculating the remainder is particularly useful as functions satisfying the conditions of Lemma 1 include the exponential, logarithm, inverse and square-root functions on any finite interval in their domains of definition.

3.5 Range bounding

A range bounder of the Chebyshev model $(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q)$ of a function $f : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ on $Y \subset Z$ is the set $[\mathcal{P}_{f,Y}^q] + \mathcal{R}_{f,Y}^q \in \mathbb{IR}$, with $[\mathcal{P}_{f,Y}^q] \supseteq \{\mathcal{P}_{f,Y}^q(\xi) \mid \xi \in [-1, 1]^n\}$. Such bounds are needed for computing binary product and univariate operations in Chebyshev model arithmetic—see Sects. 3.3 and 3.4 above, thus directly affecting the tightness of the computed Chebyshev models. Unfortunately, exact range bounding for multivariate polynomials is NP-hard, so one has to resort to some sort of over-approximation in practice.

A number of methods exist for bounding the range of multivariate polynomials in monomial form [20, 30, 36, 43], which can be readily adapted to their Chebyshev counterparts. Apart from the naive bounding of each term separately,

$$[\mathcal{P}_{f,Y}^q] = a_{Y,0}^q + \sum_{0 < |\kappa| \leq q} |a_{Y,\kappa}^q| [-1, 1], \quad (30)$$

which may produce weak bounds, this subsection presents adaptations of: (i) the methods in [33] that involves exact bounding of the first- and diagonal second-order terms; and (ii) the method of Bernstein.

3.5.1 Exact bounding of first- and diagonal second-order terms

We start by rewriting the polynomial part in the form

$$\mathcal{P}_{f,Y}^q(\xi) = \sum_{j=1}^n \left[a_{Y,\kappa^{2,j}}^q T_2(\xi_j) + a_{Y,\kappa^{1,j}}^q T_1(\xi_j) \right] + \mathcal{H}_{f,Y}^q(\xi),$$

where $\kappa^{1,j}$ and $\kappa^{2,j}$ are the multi-indices given by $\kappa_i^{1,j} := \delta_{i,j}$ and $\kappa_i^{2,j} := 2\delta_{i,j}$ for each $j \in \{1, \dots, n\}$, using the Kronecker δ notation; and $\mathcal{H}_{f,Y}^q$ is the multivariate polynomial containing the same terms as $\mathcal{P}_{f,Y}^q$ apart from those indexed by κ_j^1 and κ_j^2 . Then, a similar rearrangement to the one proposed in [33] can be obtained by using the property that $T_2(\xi) = 2T_1(\xi)^2 - 1$,

$$\mathcal{P}_{f,Y}^q(\xi) = \sum_{j=1}^n \left[2a_{Y,\kappa^{2,j}}^q \left(T_1(\xi_j) + \frac{a_{Y,\kappa^{1,j}}^q}{4a_{Y,\kappa^{2,j}}^q} \right)^2 - a_{Y,\kappa^{2,j}}^q - \frac{(a_{Y,\kappa^{1,j}}^q)^2}{8a_{Y,\kappa^{2,j}}^q} \right] + \mathcal{H}_{f,Y}^q(\xi).$$

This way, an interval enclosure is obtained as

$$[\mathcal{P}_{f,Y}^q] = \sum_{j=1}^n \left[2a_{Y,\kappa^{2,j}}^q \left([-1, 1] + \frac{a_{Y,\kappa^{1,j}}^q}{4a_{Y,\kappa^{2,j}}^q} \right)^2 - a_{Y,\kappa^{2,j}}^q - \frac{(a_{Y,\kappa^{1,j}}^q)^2}{8a_{Y,\kappa^{2,j}}^q} \right] + [\mathcal{H}_{f,Y}^q],$$

with $[\mathcal{H}_{f,Y}^q]$ computed directly from (30). This bounder may only be considered when $|a_{Y,\kappa}^{q,2,j}| \geq \epsilon$, where ϵ is a small positive number; otherwise, separate bounding of the linear and quadratic terms for the corresponding component j as in (30) may be used.

3.5.2 Bounding method of Bernstein

The Bernstein algorithm is a well-established tool for computing bounds on the range of multivariate polynomials over a box; see, e.g., [20, 30] and references therein. The procedure involves rewriting the multivariate polynomial from Chebyshev form into Bernstein form:

$$\forall \xi \in [-1, 1]^n, \quad \mathcal{P}_{f,Y}^q(\xi) = \sum_{\kappa \in \{0, \dots, r\}^n} b_{Y,\kappa}^q \prod_{i=1}^n B_{\kappa_i}^r \left(\frac{\xi_i + 1}{2} \right), \quad (31)$$

with $r \geq q$, and where B_i^r , with $i \leq r$, denotes the i th Bernstein polynomial of order r on $[0, 1]$, given by

$$B_i^r(z) = \binom{r}{i} z^i (1-z)^{r-i}.$$

In particular, the Bernstein coefficient $b_{Y,\kappa}^q$ in (31) can be expressed in terms of the Chebyshev coefficients $a_{Y,\kappa}^q$ as follows [48]

$$\forall \kappa \in \{0, \dots, r\}^n, \quad b_{Y,\kappa}^q = \sum_{|v| \leq q} a_{Y,v}^q \prod_{i=1}^n M_{\kappa_i, \gamma_i}$$

$$\text{with: } M_{k,j} := \binom{r}{k}^{-1} \sum_{i=\max\{0, j+k-r\}}^{\min\{j,k\}} (-1)^{j-i} \binom{2j}{2i} \binom{r-j}{k-i}.$$

This transformation turns out to be particularly well-conditioned from a numerical stability standpoint [48]. At this point, an interval enclosure of $\mathcal{P}_{f,Y}^q$ on $[-1, 1]^n$ is simply obtained as

$$[\mathcal{P}_{f,Y}^q] = \left[\min_{v \in \{0, \dots, r\}^n} b_{Y,v}^q, \max_{v \in \{0, \dots, r\}^n} b_{Y,v}^q \right].$$

The main advantage of this approach is that the foregoing enclosure converges to the actual polynomial range in the Hausdorff sense as the order r increases [20]. Moreover, it is possible to bound the actual under- or over-estimation for a given order $r \geq q$.

4 Convergence analysis

This section presents an analysis of the local convergence rate of Chebyshev models constructed from the application of the arithmetic operations described in Sect. 3. The global convergence of this arithmetic as the polynomial expansion order q increases is also discussed at the end of the section.

Definition 2 Let the function $f : Z \rightarrow \mathbb{R}$ be defined on $Z \subseteq \mathbb{R}^n$. The scheme $(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q)_{Y \subset Z}$ of q th-order Chebyshev models for f , with

$$\mathcal{P}_{f,Y}^q(\xi) := \sum_{|\kappa| \leq q} a_{Y,\kappa}^q T_{\kappa}(\xi),$$

is said to have *local convergence order* r on Z if

$$\forall \kappa \in \mathbb{N}^n \quad \text{with } |\kappa| \leq q, \quad a_{Y,\kappa}^q = \mathcal{O}\left((\text{rad } Y)^{\min\{|\kappa|, r\}}\right) \quad (32)$$

$$\text{and } \text{rad } \mathcal{R}_{f,Y}^q = \mathcal{O}\left((\text{rad } Y)^r\right), \quad (33)$$

for all interval vectors $Y \subset Z$ with sufficiently small $\text{rad } Y$.

Notice the extra conditions (32) imposed on the coefficients of the polynomial approximant in the previous definition. While these conditions are not strictly necessary to establish local convergence, the propagation of (local) convergence order for the polynomial coefficients through binary sum, binary product and univariate outer-composition operations provides insight into the propagation of the convergence order for the remainder term in the following analysis.

As far as initialization is concerned, it is clear that a scheme of q th-order Chebyshev models $(\mathcal{P}_{z_i,Y}^q, \mathcal{R}_{z_i,Y}^q)_{Y \subset Z}$ for the variable $z_i \in Z_i$ with $i \in \{1, \dots, n\}$, as given by (11) and (12), has local convergence order 1 if $q = 0$, and infinite order if $q \geq 1$. In both cases, the schemes of q th-order Chebyshev models for the variables thus have convergence order no less than $q + 1$. Regarding the range bounders introduced in Sect. 3.5, the convergence condition (32) imposes that $\text{rad}([\mathcal{P}_{f,Y}^q] + \mathcal{R}_{f,Y}^q) = \mathcal{O}(\text{rad } Y)$ whenever the scheme $(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q)_{Y \subset Z}$ has convergence order $r \geq 1$.

Next, we investigate the local convergence order of schemes of Chebyshev models as propagated through factorable functions, with the corresponding binary sum, binary product and univariate outer-composition operations derived in Sects. 3.2–3.4. With regards to binary sum operations given by (13)–(14), the sum of two schemes of q th-order Chebyshev models trivially preserves the least convergence order of the operands, be they smaller than, equal to, or larger than $q + 1$. In particular, the addition/subtraction of two schemes with local convergence order $q + 1$ is itself a scheme of order $q + 1$. The propagation of convergence orders through binary product and univariate outer-composition operations is somewhat more involved, and detailed in Sects. 4.1 and 4.2.

4.1 Local convergence rate of binary product operations

Adopting the notation introduced in Sect. 3.3, we assume here that the schemes of q th-order Chebyshev models $(\mathcal{P}_{f_1,Y}^q, \mathcal{R}_{f_1,Y}^q)_{Y \subset Z}$ and $(\mathcal{P}_{f_2,Y}^q, \mathcal{R}_{f_2,Y}^q)_{Y \subset Z}$ have local convergence orders $r_1 \geq 1$ and $r_2 \geq 1$ on Z , respectively.

The following lemma is instrumental to prove convergence of the Chebyshev coefficients in the product scheme:

Lemma 2 *For any $q \geq 0$, any $\kappa \in \mathbb{N}^q$, and any pair $(\lambda, \mu) \in \mathbb{P}_q(\kappa)$, we have $|\lambda + \mu| \geq |\kappa|$.*

Proof By construction of $\mathbb{P}_q(\kappa)$ in (15), we either have $\lambda_i + \mu_i = \kappa_i$ or $|\lambda_i - \mu_i| = \kappa_i$, for each $i \in \{1, \dots, n\}$. Therefore, the result follows by noting that $\lambda_i + \mu_i \geq |\lambda_i - \mu_i|$. \square

By the construction of the product polynomial in (16) and by Lemma 2, we have

$$\left| c_{Y,\kappa}^q \right| \leq \sum_{(\lambda,\mu) \in \mathbb{P}_q(\kappa)} \frac{\left| a_{Y,\lambda}^q \right| \left| b_{Y,\mu}^q \right|}{2^{N(\lambda,\mu)}} = \mathcal{O}\left((\text{rad } Y)^{\min\{r_1, r_2, |\kappa|\}}\right), \quad (34)$$

for every $\kappa \in \mathbb{N}^n$ with $|\kappa| \leq 2q$. Therefore, the product rule (17) propagates the least convergence order of the operands through the polynomial coefficients, up to order $|\kappa|$. Moreover, we have $[\mathcal{P}_{f_1,Y}^q] = \mathcal{O}(1)$, $[\mathcal{P}_{f_2,Y}^q] = \mathcal{O}(1)$, and

$$\sum_{q < |\kappa| \leq 2q} |c_{Y,\kappa}^q| [-1, 1] = \mathcal{O}\left((\text{rad } Y)^{\min\{r_1, r_2, q+1\}}\right),$$

so it follows from the remainder propagation rule (18) that

$$\text{rad } \mathcal{R}_{f_1 f_2, Y}^q = \mathcal{O}\left((\text{rad } Y)^{\min\{r_1, r_2, q+1\}}\right). \quad (35)$$

Overall, the least convergence order of the operands thus propagates through the remainder, up to order $q + 1$. In particular, the product scheme has local convergence order $q + 1$ whenever both operands have convergence order $q + 1$. These results are summarized in the following theorem:

Theorem 1 *Let $(\mathcal{P}_{f_1,Y}^q, \mathcal{R}_{f_1,Y}^q)_{Y \subset Z}$ and $(\mathcal{P}_{f_2,Y}^q, \mathcal{R}_{f_2,Y}^q)_{Y \subset Z}$ be schemes of q th-order Chebyshev models, with local convergence orders $r_1 \geq 1$ and $r_2 \geq 1$, respectively, for the functions $f_1, f_2 : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Then, the product scheme $(\mathcal{P}_{f_1 f_2, Y}^q, \mathcal{R}_{f_1 f_2, Y}^q)_{Y \subset Z}$ given by (17) and (18) has local convergence order $\min\{r_1, r_2, q + 1\}$.*

4.2 Local convergence rate of univariate outer-composition operations

Using the notation introduced in Sect. 3.4, we assume that the outer functions $F : X \subseteq \mathbb{R} \rightarrow \mathbb{R}$ is s -times continuously differentiable, and we consider the scheme of q th-order Chebyshev models $(\mathcal{P}_{F,W}^q, \mathcal{R}_{F,W}^q)_{W \subset X}$, as constructed from (25), (26) [resp. from (26), (28)] if $s \geq q + 1$; or otherwise constructed from (25), (27) [resp. from (28), (29)] if $F^{(s)}$ has a finite total variation on X .

Lemma 3 *The scheme $(\mathcal{P}_{F,W}^q, \mathcal{R}_{F,W}^q)_{W \subset X}$ has local convergence order $\min\{s, q + 1\}$ on X .*

Proof See “Appendix 2”. □

To carry out the analysis, we also assume that the scheme of q th-order Chebyshev models $(\mathcal{P}_{f,Y}^q, \mathcal{R}_{f,Y}^q)_{Y \subset Z}$ for the inner functions $f : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ has local convergence order $r \geq 1$. Notice that the Clenshaw summations (20)–(22) consists of binary sum and product operations only and is initialized with constants. However, by construction of $(\mathcal{P}_{\phi,Y}^q, \mathcal{R}_{\phi,Y}^q)_{Y \subset Z}$ in (19), we have

$$a_{Y,0}^q = 0, \quad \forall \kappa \in \mathbb{N}^n \text{ with } 1 \leq |\kappa| \leq q, \quad a_{Y,\kappa}^q = \mathcal{O}\left((\text{rad } Y)^{\min\{|\kappa|, r-1\}}\right),$$

and $\text{rad } \mathcal{R}_{\phi,Y}^q = \mathcal{O}\left((\text{rad } Y)^{r-1}\right),$

where $a_{Y,\kappa}^q$ denote the coefficients of $\mathcal{P}_{\phi,Y}^q$. Even though $(\mathcal{P}_{\phi,Y}^q, \mathcal{R}_{\phi,Y}^q)_{Y \subset Z}$ has local convergence order $r - 1$ only, the following lemma shows that the original order r of the inner operand nonetheless propagates through the scheme $(\mathcal{P}_{\mathcal{P}_{F \circ \phi, Y}}^q, \mathcal{R}_{\mathcal{P}_{F \circ \phi, Y}}^q)_{Y \subset Z}$, up to order $q + 1$:

Lemma 4 *The scheme $(\mathcal{P}_{\mathcal{P}_{F \circ \phi, Y}}^q, \mathcal{R}_{\mathcal{P}_{F \circ \phi, Y}}^q)_{Y \subset Z}$ generated through the Clenshaw recursion (20)–(22) satisfies*

$$\forall \kappa \in \mathbb{N}^n \text{ with } |\kappa| \leq q, \quad c_{Y,\kappa}^q = \mathcal{O}\left((\text{rad } Y)^{\min\{|\kappa|, r, s\}}\right),$$

$$\text{rad } \mathcal{R}_{\mathcal{P}_F \circ \Phi, Y}^q = \mathcal{O} \left((\text{rad } Y)^{\min\{r, q+1\}} \right),$$

where $c_{Y, \kappa}^q$ denote the coefficients of $\mathcal{P}_{\mathcal{P}_F \circ \Phi, Y}^q$.

Proof See “Appendix 3”. \square

A direct consequence of Lemmata 3 and 4 is that the remainder term $\mathcal{R}_{\mathcal{P}_F \circ \Phi, Y}^q$, as defined in (24), has convergence order $\min\{r, s, q+1\}$. Therefore, the convergence order of the inner operand may propagate through the remainder, up to order $q+1$, in a q th-order scheme of Chebyshev models. Moreover, if the outer operand is s -times continuously differentiable, the convergence order of the composition scheme may not be larger than s . In particular, the composition scheme has convergence order $q+1$ whenever the inner scheme has convergence order $q+1$ and the outer function is $q+1$ -times continuously differentiable. These results are summarized in the following theorem:

Theorem 2 Let $(\mathcal{P}_{F, W}^q, \mathcal{R}_{F, W}^q)_{W \subset X}$ be a scheme of q th-order Chebyshev models for the s -times continuously-differentiable function $F : X \subseteq \mathbb{R} \rightarrow \mathbb{R}$, as constructed from (25), (26) [resp. from (26), (28)] if $s \geq q+1$; or otherwise constructed from (25), (27) [resp. from (28), (29)] if $F^{(s)}$ has a finite total variation on X . Let also $(\mathcal{P}_{f, Y}^q, \mathcal{R}_{f, Y}^q)_{Y \subset Z}$ be a scheme of q th-order Chebyshev models for the function $f : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ with local convergence order $r \geq 1$. Then, the composition scheme $(\mathcal{P}_{F \circ f, Y}^q, \mathcal{R}_{F \circ f, Y}^q)_{Y \subset Z}$ given by (23) and (24) has convergence order $\min\{r, s, q+1\}$ on Z .

4.3 Global convergence of Chebyshev model arithmetic

The previous convergence analysis investigates the rate at which the remainder bound shrinks as the domain Y of the variables is progressively reduced to a point, for a given expansion order q . This section discusses a different type of convergence, namely the property of the remainder bound to shrink to zero upon increasing the expansion order $q \rightarrow \infty$, for a fixed variable domain Y .

Definition 3 Let the function $f : Z \rightarrow \mathbb{R}$ be defined on $Z \subseteq \mathbb{R}^n$, and let $\mathbb{R}^n \ni Y \subset Z$. The scheme $(\mathcal{P}_{f, Y}^q, \mathcal{R}_{f, Y}^q)_{q \geq 0}$ of q th-order Chebyshev models for f , with

$$\mathcal{P}_{f, Y}^q(\xi) := \sum_{|\kappa| \leq q} a_{Y, \kappa}^q T_{\kappa}(\xi),$$

is said to be *globally convergent* on Y if

$$\lim_{q \rightarrow \infty} \sum_{|\kappa| \leq q} |a_{Y, \kappa}^q| = \bar{A}_Y < \infty \quad \text{and} \quad \lim_{q \rightarrow \infty} \mathcal{R}_{f, Y}^q = \{0\}. \quad (36)$$

As far as initialization is concerned, it is clear that a scheme of q th-order Chebyshev models $(\mathcal{P}_{y_i, Y}^q, \mathcal{R}_{y_i, Y}^q)_{q \geq 0}$ for the variable $y_i \in Y_i$ with $i \in \{1, \dots, n\}$, as given by (11) and (12), is globally convergent on Y . Regarding the range bounds introduced in Sect. 3.5, it is also worth noting that the sequence $\{\text{rad } [\mathcal{P}_{f, Y}^q]\}_{q \geq 0}$ is bounded when (36) holds.

Next, we adopt the notation introduced in Sects. 3.2 and 3.3, and we assume that the schemes $(\mathcal{P}_{f_1, Y}^q, \mathcal{R}_{f_1, Y}^q)_{q \geq 0}$ and $(\mathcal{P}_{f_2, Y}^q, \mathcal{R}_{f_2, Y}^q)_{q \geq 0}$ are globally convergent on Y . The summation/subtraction scheme $(\mathcal{P}_{f_1 \pm f_2, Y}^q, \mathcal{R}_{f_1 \pm f_2, Y}^q)_{q \geq 0}$, as given in (14), is trivially globally convergent by the triangle inequality. Concerning the product operation in Sect. 3.3, we have from (16) that

$$\begin{aligned}
\sum_{|\kappa| \leq q} |c_{Y,\kappa}^q| &= \sum_{|\kappa| \leq q} \left| \sum_{(\lambda, \mu) \in \mathbb{P}_q(\kappa)} \frac{a_{Y,\lambda}^q b_{Y,\mu}^q}{2^{N(\lambda, \mu)}} \right| \\
&\leq \sum_{|\kappa| \leq q} \sum_{(\lambda, \mu) \in \mathbb{P}_q(\kappa)} \frac{|a_{Y,\lambda}^q| |b_{Y,\mu}^q|}{2^{N(\lambda, \mu)}} \\
&= \sum_{|\lambda| \leq q} \sum_{|\mu| \leq q} |a_{Y,\lambda}^q| |b_{Y,\mu}^q|,
\end{aligned}$$

so

$$\lim_{q \rightarrow \infty} \sum_{|\kappa| \leq q} |c_{Y,\kappa}^q| \leq \lim_{q \rightarrow \infty} \sum_{|\lambda| \leq q} |a_{Y,\lambda}^q| \sum_{|\mu| \leq q} |b_{Y,\mu}^q| < \infty,$$

and, in particular,

$$\lim_{q \rightarrow \infty} \sum_{q < |\kappa| \leq 2q} |c_{Y,\kappa}^q| \leq \lim_{q \rightarrow \infty} \sum_{|\kappa| > q} |c_{Y,\kappa}^q| = 0.$$

It then follows from (18) that the product scheme $(\mathcal{P}_{f_1 f_2, Y}^q, \mathcal{R}_{f_1 f_2, Y}^q)_{q \geq 0}$ is globally convergent on Y , as long as the sequences $\{\text{rad}[\mathcal{P}_{f_1, Y}^q]\}_{q \geq 0}$ and $\{\text{rad}[\mathcal{P}_{f_2, Y}^q]\}_{q \geq 0}$ are bounded. These considerations are summarized in the following theorem:

Theorem 3 *Let $(\mathcal{P}_{f_1, Y}^q, \mathcal{R}_{f_1, Y}^q)_{Y \subset Z}$ and $(\mathcal{P}_{f_2, Y}^q, \mathcal{R}_{f_2, Y}^q)_{Y \subset Z}$ be globally convergent schemes of Chebyshev models for the functions $f_1, f_2 : Z \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, respectively. Then, the summation/subtraction scheme $(\mathcal{P}_{f_1 \pm f_2, Y}^q, \mathcal{R}_{f_1 \pm f_2, Y}^q)_{q \geq 0}$ given by (13) and (14) is globally convergent on Y ; and so is the product scheme $(\mathcal{P}_{f_1 f_2, Y}^q, \mathcal{R}_{f_1 f_2, Y}^q)_{q \geq 0}$ given by (17) and (18) for the range bounding methods in Sect. 3.5.*

In order to formally establish global convergence of the Chebyshev model arithmetic presented in Sect. 3, one also needs to establish sufficient conditions on both the inner and outer operands of a composition operation in order for the resulting composition scheme to be globally convergent. The development of such conditions calls for an analysis of the propagation of the global convergence property through the Clenshaw recurrence, for which results are currently lacking. The numerical case studies presented in the following sections suggest that global convergence might be guaranteed under mild regularity conditions nonetheless, and this will be the topic of further research.

5 Numerical implementation and case studies

This section presents numerical case studies that illustrate the convergence properties of Chebyshev models, and makes comparisons with Taylor models in terms of tightness and computational effort. All the computations that led to these results are performed with the library MC++ (<http://omega-icl.bitbucket.org/mcpp/>), which features classes for the evaluation of factorable functions in Taylor and Chebyshev arithmetics (along with other arithmetics). Verified interval libraries, such as PROFIL (<http://www.ti3.tu-harburg.de/>) or FILIB++ (<http://www2.math.uni-wuppertal.de/~xsc/>), are used to perform the remainder interval computations, but any round-off caused by operations between the polynomial coefficients is not accounted for in the current implementation. We note that approaches to account

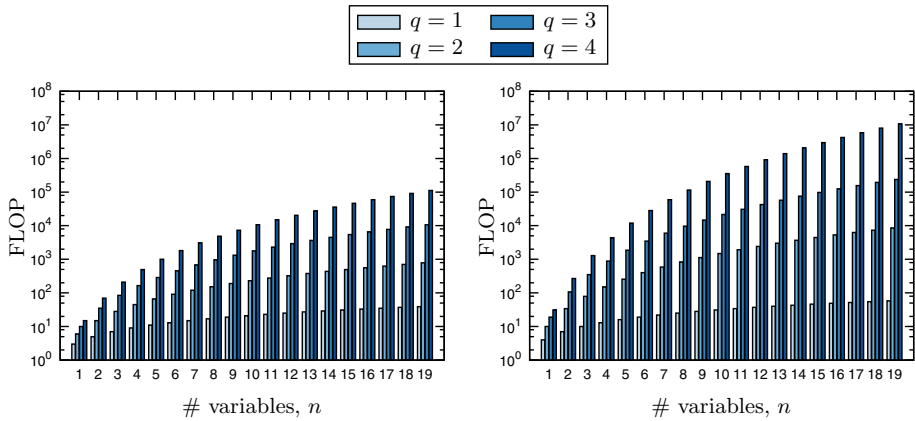


Fig. 1 Number of FLOP for performing polynomial product in monomial basis (*left plot*) and in Chebyshev basis (*right plot*)

for errors due to floating-point arithmetic are well-documented for Taylor models [36,43], and the very same approaches may be used for Chebyshev models in order to arrive at a fully verified implementation.

MC++ implements the Chebyshev model arithmetic and range bounding operations described in Sect. 3, based on a dense representation of the multivariate polynomials. If not otherwise noted, Chebyshev models for the intrinsic functions are computed based on the Chebyshev interpolation polynomial. Moreover, the exact remainder bounding approach established in Sect. 3.4 is used for the intrinsic functions $\exp(x)$, $\log(x)$, \sqrt{x} , and $\frac{1}{x}$, and the default range binder uses exact bounding of first- and diagonal second-order terms as described in Sect. 3.5. Another improvement involves intersecting the polynomial model bounds with those derived from simple interval analysis at each operation as a means of tightening the remainder bound, avoiding division by zero, etc; this extension is similar in essence to the mixed affine-arithmetic/interval-arithmetic model by Stolfi and Figueiredo [19].

The most critical operation in an efficient implementation is multiplication of Chebyshev models, since it is used extensively in univariate outer-composition operations besides bivariate product operations. The approach used in MC++ involves constructing the index sets $\mathbb{P}_q(\kappa)$ given by (15), for all $\kappa \in \mathbb{N}^n$ with $|\kappa| \leq q$, prior to propagating the Chebyshev models through a DAG of the factorable function. The number of floating-point operations (FLOP) required to compute the product polynomial $\mathcal{P}_{f_1 f_2, Y}^q$ in (17) is shown on the left plot of Fig. 1 for Chebyshev expansions of order $q = 1, \dots, 4$ and for functions with up to $n = 19$ variables. Even for low expansion orders does performing polynomial products in Chebyshev basis result in an increase by 1 or 2 orders of magnitude in the number of FLOP, compared with monomial basis which is shown on the right plot. The better approximation capability of Chebyshev models may thus come at the price of a much higher computational burden compared with Taylor models, at least in a dense implementation. This refinement calls for the development of sparse implementations for both Taylor and Chebyshev model arithmetics in order to overcome this limitation and tackle larger-scale problems, a topic for future research.

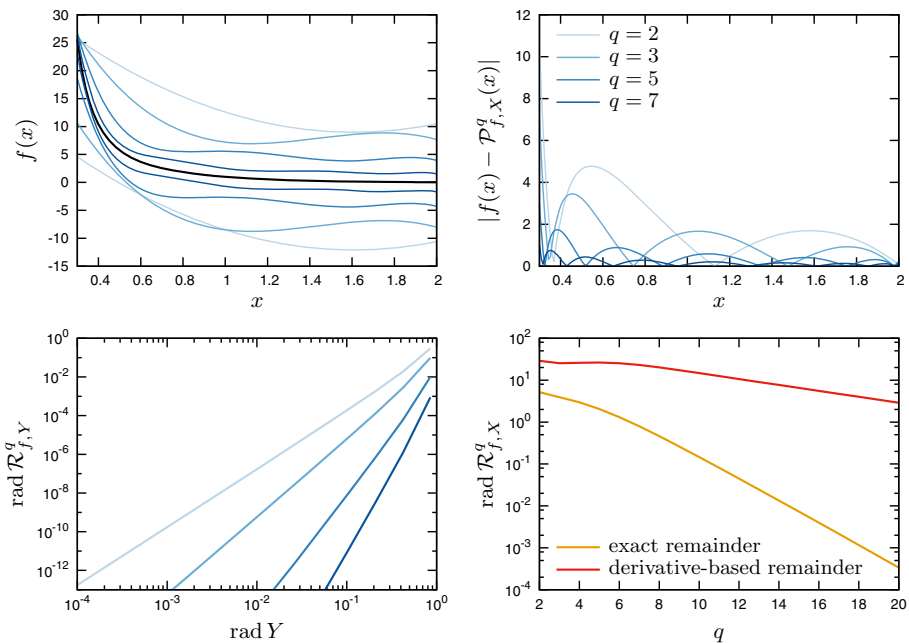


Fig. 2 Application of Chebyshev models to Example 1. *Top-left plot* polynomial model enclosures of the function f (black line) for expansion orders $q = 2, 3, 5, 7$. *Top-right plot* approximation error for expansion orders $q = 2, 3, 5, 7$. *Bottom-left plot* local convergence analysis for a sequence of shrinking interval sets $Y \subset X$ around $x = 1$. *Bottom-right plot* global convergence analysis on X as $q \rightarrow \infty$ with the exact remainder estimate compared with the high-order derivative-based remainder formula

5.1 Example 1

We consider the univariate function f given by

$$f(x) = \exp\left(-x^2 + \frac{1}{x}\right),$$

for the variable $x \in X := [0.3, 2]$. A number of Chebyshev model enclosures for various expansion orders q in the range 2–7 are shown on the top-left plot of Fig. 2, along with the corresponding approximation errors on the opposite plot. The scheme of Chebyshev models appears to be globally convergent as $q \rightarrow \infty$, both with the derivative-based remainder formula and with the exact remainder formula for the univariate outer-composition operations, as shown on the bottom-right plot of Fig. 2. However, the use of the exact remainder significantly tightens the bounds and improves the rate of convergence. In contrast, classical Taylor models fail to converge to the function as the order q is increased, despite the function f being analytic on the set X . This is because the variable range of interest is partially outside the radius of convergence for this function, which also means that the width of the interval remainder increases with q . Finally, the bottom-left plot on Fig. 2 shows that local convergence of q th-order Chebyshev model is of order $q + 1$ around $x = 1$. This observation is consistent with the analysis conducted in Sect. 4 since f is analytic.

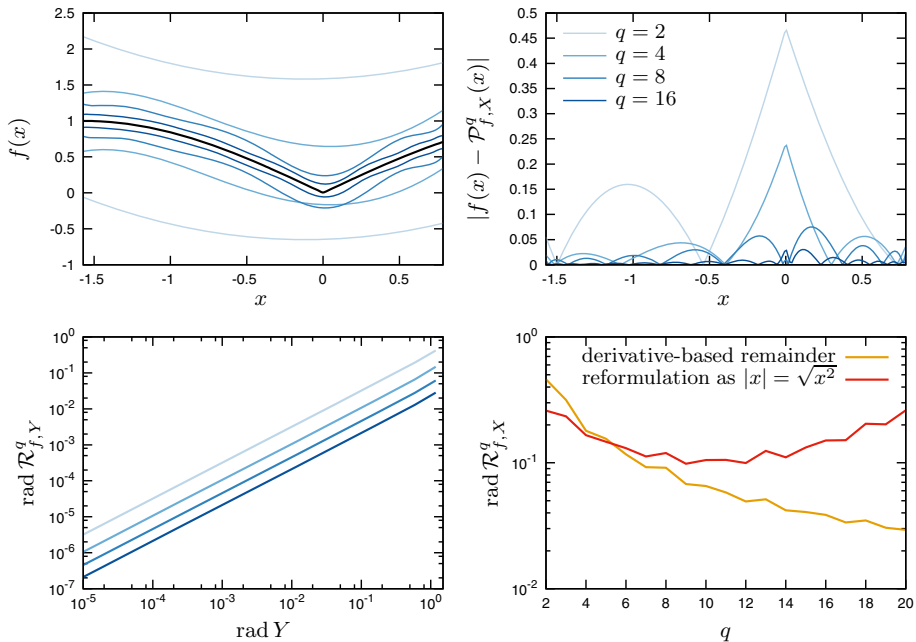


Fig. 3 Application of Chebyshev models to Example 2. *Top-left plot* Polynomial model enclosures of the function f (black line) for expansion orders $q = 2, 4, 8, 16$. *Top-right plot* approximation error for expansion orders $q = 2, 4, 8, 16$. *Bottom-left plot* local convergence analysis for a sequence of shrinking interval sets $Y \subset X$ around $x = 0$. *Bottom-right plot* global convergence analysis on X as $q \rightarrow \infty$ with the derivative-based remainder formula or the reformulation as $|x| = \sqrt{x^2}$

5.2 Example 2

We consider the univariate function f given by

$$f(x) = \sin |x|,$$

for the variable $x \in X := [-\frac{\pi}{2}, \frac{\pi}{4}]$. Notice that classical Taylor models may not be used to bound the function f on X given the nonsmoothness at $x = 0$. In contrast, Chebyshev models can be constructed since the univariate $\sin(x)$ is analytic in \mathbb{R} , whereas the derivative of $|x|$ has total bounded variation on X . The top plots in Fig. 3 shows how the bounds improve with a larger expansion order q , up to $q = 16$. It is evident that a much higher-order expansion is necessary in comparison with the smooth function in the previous example (Sect. 5.1), and that the approximation error is the largest close to $x = 0$. In agreement with the theory in Sect. 4, the local convergence of any q th-order Chebyshev model at this point is of order 1, regardless of the expansion order $q \geq 0$; a prediction that is confirmed in the bottom-left plot of Fig. 3. But despite the nonsmoothness at $x = 0$ and the linear local convergence around this point, Chebyshev models are found to be globally convergent on X as $q \rightarrow \infty$, when the remainder formula (29) is used for bounding the remainder of the univariate term $|x|$ (bottom-right plot on Fig. 2).

Notice that this function may also be reformulated by substitution of $|x| = \sqrt{x^2}$, thus providing an alternative route to creating the Chebyshev model. This reformulation makes the use of the exact remainder estimate for the square-root term possible, which explains

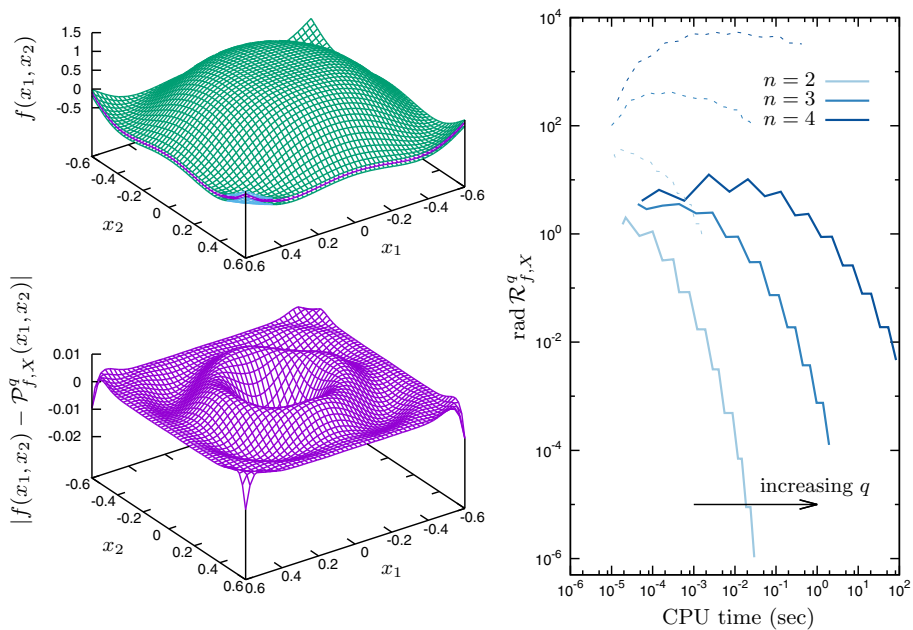


Fig. 4 Application of Chebyshev models to Example 3. *Top-left plot* polynomial model enclosing the function f for $n = 2$ with expansion orders $q = 8$. *Bottom-left plot* corresponding approximation error for expansion orders $q = 8$. *Right plot* computational effort and global convergence of the Chebyshev remainder $\text{rad } \mathcal{R}_{f,X}^q$ with expansion orders $q = 2, \dots, 20$ for dimensions $n = 2$ to 4 (solid lines), and comparison with their Taylor models counterparts (dashed lines)

why tighter bounds for Chebyshev models of order $q = 5$ or less are obtained here with the reformulation approach (bottom-right plot of Fig. 3). However, the reformulation also causes Chebyshev models to diverge on X as $q \rightarrow \infty$, a behavior attributed to the fact that the total variation of the square-root term is unbounded on intervals containing zero.

5.3 Example 3

As a final example, we consider the multivariate function f given by

$$f(x) = \exp\left(\sum_{i=1}^n x_i\right) \cos\left(2\pi \sum_{i=1}^n x_i\right),$$

for the variable $x \in X := [-0.6, 0.6]^n$, with the dimension n taken as 2, 3 or 4. The Chebyshev model enclosure and corresponding approximation error in the 2-variable case are presented on the left part of Fig. 4 for an expansion of order $q = 8$, showing a tight approximation of the function. From the right plot of Fig. 4, it is seen that the computational burden increases with both the expansion order q and the dimension n of the problem. Moreover, computing Chebyshev models (solid lines) is more demanding than Taylor models (dashed lines), due to the extra complexity of the polynomial product in Chebyshev basis (see Fig. 1 above and related discussion). However, this extra complexity pays off in terms of the tightness of the resulting bounds. For $n = 2$, one must consider expansions of order greater than $q = 18$ in order for a Taylor model's remainder bound to be as tight as those given by

a second-order Chebyshev models only. For $n = 3$ or $n = 4$ likewise, although the Taylor models appear to be globally convergent on X , an expansion order far greater than $q = 20$ would be necessary in order to be competitive with just second-order Chebyshev models. Overall, Chebyshev models thus outperform their Taylor counterparts. Finally, it is interesting to note that as the function f being even, odd ordered polynomial models prove to provide tighter bounds than even ones, thus explaining why there is sometimes little improvement (or even deterioration) in the bounds even though the computational time increases.

6 Conclusions

This paper has formalized an arithmetic for the propagation of Chebyshev models through binary sums, binary products and univariate outer-composition operations in factorable functions. A simple method to calculate the exact remainder bounds for certain univariate terms, including $\exp(\cdot)$, $\log(\cdot)$, $\sqrt{\cdot}$ and $(\cdot)^{-1}$, is established. Adaptations of existing range bounders for Taylor models in order to use them with Chebyshev models are also discussed, which are a vital part of polynomial model arithmetic. The local convergence of Chebyshev models has been analyzed and proven to be equivalent to Taylor models, although the Chebyshev model remainder bounds are often found to be orders of magnitude better than their Taylor model counterparts over larger variable domains. This behavior is supported by the result that Chebyshev models are globally convergent through addition/subtraction and multiplication operations. The global convergence property is also conjectured for composition of Chebyshev models under mild conditions, although sufficient conditions are yet to be formally established. Such convergence properties have been illustrated through several numerical examples, based on an implementation in the authors' in-house library MC++. These examples also illustrate some of the advantages of Chebyshev models over Taylor models, including being able to bound functions with points of non-smoothness or a divergent Taylor expansion. Even though Chebyshev models of equivalent order are computationally more expensive to create than their Taylor counterparts, mainly due to the much greater number of operations required for the binary product operation, they provide benefits by being significantly tighter. This may allow for a much lower order polynomial model to be used and creates a net benefit in terms of computational effort.

On the computational side, the current implementation creates a polynomial which is of the same order for all the variables and is stored using a dense representation. For many applications however, a large number of coefficients may be equal to zero, thereby calling for a sparse implementation of Chebyshev model arithmetic. Such an implementation would reduce both the computational time and the memory requirements. Another area of improvement would concern the binary product operation, whose complexity currently scales exponentially in the number of variables and the expansion order. Several studies have explored approaches to speeding up the multiplication of univariate and bivariate polynomials in Chebyshev basis (e.g., [2, 5, 22, 45]). Ways to use some of these developments for multivariate Chebyshev polynomials could be explored as part of future work. Binary product operations are also where much of the overestimation occurs, given that the polynomial part of the Chebyshev model product may not match the Chebyshev expansion of the product function. Instead, it might be possible to directly create multivariable Chebyshev approximations, for instance by following a similar approach as in `chebfun2` [56]. This idea is also similar to the recent work on multivariate McCormick relaxations [59], where instead of decomposing the factorable function down to binary sums, products and univariate compositions, it would be possible to directly bound some multivariate terms.

Acknowledgments Financial support from Marie Curie Career Integration Grant PCIG09-GA-2011-293953 and the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/J006572/1 is gratefully acknowledged. JR thanks EPSRC and the Department of Chemical Engineering at Imperial College London for doctoral training awards. MEV thanks CONACYT for doctoral scholarship.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Proof of Lemma 1

Since F is Lipschitz continuous on $[-1, 1]$, it has a unique representation as an absolutely and uniformly convergent Chebyshev series, and we have

$$\forall \xi \in [-1, 1], \quad F(\xi) - \mathcal{P}_F^q(\xi) = \sum_{k=q+1}^{\infty} a_k T_k(\xi),$$

with the Chebyshev coefficients given by

$$a_k := \frac{2}{\pi} \int_0^\pi f(\cos(\theta)) \cos(k\theta) d\theta. \quad (37)$$

Observe that $\|F - \mathcal{P}_F^q\| = \sum_{k=q+1}^{\infty} |a_k|$ under the following scenarios:

- (i) the coefficients a_k with $k \geq q+1$ are of the same sign, either non-negative or non-positive, in which case $|F(1) - \mathcal{P}_F^q(1)| = \sum_{k=q+1}^{\infty} |a_k|$ by the property that $T_k(1) = 1, \forall k \geq 0$;
- (ii) the coefficients a_k with $k \geq q+1$ alternate sign, either non-negative for even coefficients and non-positive for odd coefficients or vice-versa, in which case $|F(-1) - \mathcal{P}_F^q(-1)| = \sum_{k=q+1}^{\infty} |a_k|$ by the property that $T_k(-1) = (-1)^k, \forall k \geq 0$.

Without loss of generality, we shall prove that case (i) holds with non-negative Chebyshev coefficients a_k when F has an analytic extension on the closed unit disk and all its successive derivatives are nonnegative on $[-1, 1]$; the other three cases follow readily by symmetry, considering $-F(\xi)$, $F(-\xi)$ or $-F(-\xi)$.

Under the assumption that F has an analytic extension on the closed unit disk, its Taylor series at any point $\xi \in [-1, 1]$ converges to F on $[-1, 1]$. At $\xi = 0$ in particular, we have

$$\forall \xi \in [-1, 1], \quad F(\xi) = \sum_{k=0}^{\infty} \frac{F^{(k)}(0)}{k!} \xi^k = \sum_{k=0}^{\infty} a_k T_k(\xi^k). \quad (38)$$

Recall that any polynomial \mathcal{P}^q of order $q \geq 1$, given in monomial basis by $\mathcal{P}^q(\xi) := \sum_{k=0}^q \alpha_k \xi^k$, has a unique representation in Chebyshev basis as $\mathcal{P}^q(\xi) = \sum_{k=0}^q \beta_k T_k(\xi)$ with

$$\forall k \in \{1, \dots, q\}, \quad \beta_k := \sum_{i=1}^q \frac{\alpha_i}{2^i} \sum_{j=0}^i \binom{i}{j} \delta_{k, |i-2j|}.$$

By passing to the limit $q \rightarrow \infty$, we obtain from (38) that

$$\forall k \geq 1, \quad a_k := \sum_{i=1}^{\infty} \frac{F^{(i)}(0)}{2^i i!} \sum_{j=0}^i \binom{i}{j} \delta_{k, |i-2j|}. \quad (39)$$

If follows from the nonnegativity of all the successive derivatives $F^{(i)}$ that $a_k \geq 0$ for every $k \geq 1$.

Regarding Chebyshev interpolation, on the other hand, we have

$$\forall \xi \in [-1, 1], \quad F(\xi) - \widehat{\mathcal{P}}_F^q(\xi) = \sum_{k=0}^q (a_k - \widehat{a}_k) T_k(\xi) + \sum_{k=q+1}^{\infty} a_k T_k(\xi),$$

with the coefficients \widehat{a}_k given by

$$\widehat{a}_k := \frac{2}{q+1} \sum_{j=0}^q F(\zeta_j) T_k(\zeta_j) \quad \text{with } \zeta_j := \cos\left(\frac{\pi}{2} \frac{j+1}{q+1}\right). \quad (40)$$

Like previously, we shall prove that

$$\|F - \widehat{\mathcal{P}}_F^q\| = |F(1) - \widehat{\mathcal{P}}_F^q(1)| = \pm (a_0 - \widehat{a}_0) + \sum_{k=1}^q (a_k - \widehat{a}_k) + \sum_{k=q+1}^{\infty} a_k$$

in the case that F has an analytic extension on the closed unit disk and all its successive derivatives are nonnegative on $[-1, 1]$ —the other cases following by symmetry. Since $a_k \geq 0$ for every $k \geq 1$ under these assumptions, it is sufficient to prove that $a_k \geq \widehat{a}_k$ for each $k = 1, \dots, q$. By the aliasing property of the Chebyshev polynomials [37], we have

$$\forall k \in \{0, \dots, q\}, \quad \widehat{a}_k = a_k - a_{2q+2-k} - a_{2q+2+k} + a_{4q+4-k} + a_{4q+4+k} - \dots \quad (41)$$

and, therefore, the result follows by showing that $a_k \geq a_{k+2}$ for every $k \geq 1$. Considering the case of odd coefficients first, (39) gives

$$\begin{aligned} a_{2k-1} &= \sum_{i=k}^{\infty} \frac{F^{(2i-1)}(0)}{2^{2i-2}(2i-1)!} \binom{2i-1}{i-k} \\ &= \frac{F^{(2k-1)}(0)}{2^{2k-2}(2k-1)!} + \sum_{i=k+1}^{\infty} \frac{F^{(2i-1)}(0)}{2^{2i-2}(2i-1)!} \binom{2i-1}{i-k} \\ &= \frac{F^{(2k-1)}(0)}{2^{2k-2}(2k-1)!} + \sum_{i=k+1}^{\infty} \frac{F^{(2i-1)}(0)}{2^{2i-2}(2i-1)!} \binom{2i-1}{i-k-1} \frac{i+k}{i-k} \\ &\geq \frac{F^{(2k-1)}(0)}{2^{2k-2}(2k-1)!} + a_{2k+1}, \end{aligned}$$

and a similar result holds for even coefficients. \square

Appendix 2: Proof of Lemma 3

Let the Chebyshev coefficients of the function F on the subset $W \subset X$ be denoted as

$$\forall k \geq 1, \quad a_{W,k} := \frac{2}{\pi} \int_{-1}^1 \frac{F(\text{mid } W + \xi \text{ rad } W) T_k(\xi)}{\sqrt{1-\xi^2}} d\xi.$$

Since F is s -times continuously differentiable, Taylor's theorem asserts that, for all $\xi \in [-1, 1]$, there exists $\zeta_\xi \in [-1, 1]$ such that

$$F(\text{mid } W + \xi \text{ rad } W) = \sum_{i=0}^{s-1} \frac{F^{(i)}(\text{mid } W)}{i!} (\xi \text{ rad } W)^i + \frac{F^{(s)}(\zeta_\xi)}{s!} (\xi \text{ rad } W)^s.$$

Therefore, letting $M_i := \|F^{(i)}\|_X$ for $0 \leq i \leq s$, we have

$$|a_{W,k}| \leq \sum_{i=0}^{s-1} \frac{2(\text{rad } W)^i M_i}{\pi i!} \left| \int_{-1}^1 \frac{\xi^i T_k(\xi)}{\sqrt{1-\xi^2}} d\xi \right| + \frac{2(\text{rad } W)^s M_s}{\pi s!} \int_{-1}^1 \frac{|\xi^s T_k(\xi)|}{\sqrt{1-\xi^2}} d\xi.$$

Since $\int_{-1}^1 \frac{\xi^i T_k(\xi)}{\sqrt{1-\xi^2}} d\xi = 0$ for all $i < k$, $|\int_{-1}^1 \frac{\xi^i T_k(\xi)}{\sqrt{1-\xi^2}} d\xi| \leq \pi$ for all $i \geq 0$, and $\int_{-1}^1 \frac{|\xi^s T_k(\xi)|}{\sqrt{1-\xi^2}} d\xi \leq \pi$, we obtain

$$\begin{aligned} \forall k \geq 1, \quad |a_{W,k}| &\leq (\text{rad } W)^k \sum_{i=k}^s \frac{2(\text{rad } W)^{i-k} M_i}{i!} + (\text{rad } W)^s \frac{2M_s}{s!} \\ &= \mathcal{O}\left((\text{rad } W)^{\min\{k,s\}}\right). \end{aligned}$$

Using Chebyshev interpolation likewise, a direct consequence of the aliasing property (41) is that

$$\forall k \in \{1, \dots, q\}, \quad \widehat{a}_{W,k} = \mathcal{O}\left((\text{rad } W)^{\min\{k,s\}}\right).$$

The result of the lemma follows by noting that $\text{rad } \mathcal{R}_{F,W}^q = \mathcal{O}((\text{rad } W)^{\min\{s,q+1\}})$ by (26). \square

Appendix 3: Proof of Lemma 4

We start by showing, using finite recursion on $k = q+2, q+1, \dots, 1$, that the intermediate schemes $(\mathcal{P}_{\beta_k,Y}^q, \mathcal{R}_{\beta_k,Y}^q)_{Y \subset Z}$ all satisfy

$$\forall \kappa \in \mathbb{N}^n, \quad b_{k,Y,\kappa}^q = \begin{cases} \mathcal{O}\left((\text{rad } Y)^{\min\{|\kappa|+k,r,s\}}\right) & \text{if } 0 \leq |\kappa| \leq q-k \\ 0 & \text{otherwise} \end{cases} \quad (42)$$

$$\text{rad } \mathcal{R}_{\beta_k,Y}^q = \mathcal{O}\left((\text{rad } Y)^{\min\{r,q+1\}}\right), \quad (43)$$

where $b_{k,Y,\kappa}^q$ denote the coefficients of $\mathcal{P}_{\beta_k,Y}^q$.

- This property is trivially satisfied for the schemes $(\mathcal{P}_{\beta_k,Y}^{q+1}, \mathcal{R}_{\beta_k,Y}^{q+1})_{Y \subset Z}$ and $(\mathcal{P}_{\beta_k,Y}^{q+2}, \mathcal{R}_{\beta_k,Y}^{q+2})_{Y \subset Z}$.
- Next, assume that the property holds for both $(\mathcal{P}_{\beta_k,Y}^{k+1}, \mathcal{R}_{\beta_k,Y}^{k+1})_{Y \subset Z}$ and $(\mathcal{P}_{\beta_k,Y}^{k+2}, \mathcal{R}_{\beta_k,Y}^{k+2})_{Y \subset Z}$, and consider the resulting scheme $(\mathcal{P}_{\beta_k,Y}^k, \mathcal{R}_{\beta_k,Y}^k)_{Y \subset Z}$ through (21). Since $[\mathcal{P}_{\Phi,Y}^q] = \mathcal{O}(1)$ and $[\mathcal{P}_{\beta_k,Y}^{k+1}] = \mathcal{O}((\text{rad } Y)^{\min\{k+1,r,s\}})$, the remainder of the product scheme $(\mathcal{P}_{\Phi,Y}^q, \mathcal{R}_{\Phi,Y}^q)(\mathcal{P}_{\beta_{k+1},Y}^q, \mathcal{R}_{\beta_{k+1},Y}^q)$ has order $\mathcal{O}((\text{rad } Y)^{\min\{r,q+1\}})$, and so $\mathcal{R}_{\beta_k,Y}^k$ satisfies condition (43). Concerning the polynomial part, the effect of multiplying $\mathcal{P}_{\beta_{k+1},Y}^q$ with $\mathcal{P}_{\Phi,Y}^q$ is a reduction of the order of all the coefficients $b_{k+1,Y,\kappa}^q$ by one order. Moreover,

subtracting $\mathcal{P}_{\beta_{k+2}, Y}^q$ does not change the order of the coefficients and by Lemma 3 the lead coefficient $\varphi_{B_{Y,k}^q}^q$ has order $\mathcal{O}((\text{rad } Y)^{\min\{k,s\}})$. Therefore, the coefficients $b_{k,Y,\kappa}^q$ satisfy condition (42), too.

The same reasoning applies in taking the recurrence one step further through (20), thereby giving the result. \square

References

- Adjiman, C.S., Dallwig, S., Floudas, C.A., Neumaier, A.: A global optimization method, α BB, for general twice-differentiable constrained NLPs—I. Theoretical advances. *Comput. Chem. Eng.* **22**(9), 1137–1158 (1998)
- Akleyek, S., Cenk, M., Ozbudak, F.: On the polynomial multiplication in Chebyshev form. *IEEE Trans. Comput.* **61**(4), 584–587 (2012)
- Akrotirianakis, I.G., Floudas, C.A.: A new class of improved convex underestimators for twice continuously differentiable constrained NLPs. *J. Glob. Optim.* **30**(4), 367–390 (2004)
- Armellin, R., Di Lizia, P., Makino, K., Berz, M.: Rigorous global optimization of impulsive planet-to-planet transfers in the patched-conics approximation. *Eng. Optim.* **44**(2), 133–155 (2012)
- Baszenski, G., Tasche, M.: Fast polynomial multiplication and convolutions related to the discrete cosine transform. *Linear Algebra Appl.* **252**(1), 1–25 (1997)
- Battles, Z., Trefethen, L.N.: An extension of MATLAB to continuous functions and operators. *SIAM J. Sci. Comput.* **25**, 1743–1770 (2004)
- Berz, M.: From Taylor series to Taylor models. In: *AIP Conference Proceedings*, vol. 405, pp. 1–23 (1997). doi:[10.1063/1.53493](https://doi.org/10.1063/1.53493)
- Berz, M., Hoffstätter, G.: Computation and application of Taylor polynomials with remainder bounds. *Reliab. Comput.* **4**, 83–97 (1998)
- Berz, M., Makino, K.: Rigorous global search using Taylor models. In: *Proceedings of the Conference on Symbolic Numeric Computation (SNC '2009)*, pp. 11–19, Kyoto, Japan (2009)
- Bompadre, A., Mitsos, A.: Convergence rate of McCormick relaxations. *J. Glob. Optim.* **52**(1), 1–28 (2012)
- Bompadre, A., Mitsos, A., Chachuat, B.: Convergence analysis of Taylor and McCormick–Taylor models. *J. Glob. Optim.* **57**(1), 75–114 (2013)
- Brisebarre, N., Joldeş, M.: Chebyshev interpolation polynomial-based tools for rigorous computing. In: *Proceedings of the International Symposium on Symbolic and Algebraic Computation (ISSAC '2010)*, pp. 147–154, Munich, Germany (2010)
- Chachuat, B., Houska, B., Paulen, R., Perić, N., Rajyaguru, J., Villanueva, M.E.: Set-theoretic approaches in analysis, estimation and control of nonlinear systems. *IFAC Pap. OnLine* **48**(8), 981–995 (2015)
- Clenshaw, C.W.: A note on the summation of Chebyshev series. *Math. Comput.* **9**, 118–120 (1955)
- Du, K.S., Kearfoot, R.B.: The cluster problem in multivariate global optimization. *J. Glob. Optim.* **5**(3), 253–265 (1994)
- Dzetkulić, T.: Rigorous integration of non-linear ordinary differential equations in Chebyshev basis. *Numer. Algorithms* **69**(1), 183–205 (2015)
- Eckmann, J.P., Koch, H., Wittwer, P.: A computer-assisted proof of universality in area-preserving maps. *Mem. AMS* **47**, 289 (1984)
- Elliott, D., Paget, D., Phillips, G., Taylor, P.: Error of truncated Chebyshev series and other near minimax polynomial approximations. *J. Approx. Theory* **50**(1), 49–57 (1987)
- de Figueiredo, L.H., Stolfi, J.: Affine arithmetic: concepts and applications. *Numer. Algorithms* **37**(1–4), 147–158 (2004)
- Garloff, J.: The Bernstein algorithm. *Interval Comput.* **2**, 154–168 (1993)
- Gil, A., Segura, J., Temme, N.M.: *Numerical Methods for Special Functions*. SIAM, Philadelphia (2007)
- Giorgi, P.: On polynomial multiplication in Chebyshev basis. *IEEE Trans. Comput.* **61**(6), 780–789 (2012)
- Gounaris, C.E., Floudas, C.A.: Tight convex underestimators for \mathcal{C}^2 continuous problems: II. Multivariate functions. *J. Glob. Optim.* **42**(1), 69–89 (2008)
- Hoefkens, J.: *Rigorous Numerical Analysis with High-Order Taylor Models*. Ph.D. thesis, Michigan State University, East Lansing, MI, USA (2001)
- Joldeş, M.: *Rigorous Polynomial Approximations and Applications*. Ph.D. thesis, École Normale Supérieure de Lyon, France (2011)

26. Kaucher, E., Miranker, W.L.: Self-Validating Numerics for Function Space Problems: Computation with Guarantees for Differential and Integral Equations. Academic Press, Orlando (1984)
27. Kaucher, E., Miranker, W.L.: Validated computation in a function space. Reliability in Computing, pp. 403–420. Academic Press, San Diego (1988)
28. Kurzhanski, A., Valyi, I.: Ellipsoidal Calculus for Estimation and Control. Series in Systems & Control: Foundations & Applications. Birkhäuser Basel (1997)
29. Lasserre, J.B.: Moments. Positive Polynomials and Their Applications. Imperial College Press, London (2009)
30. Lin, Q., Rokne, J.G.: Methods for bounding the range of a polynomial. J. Comput. Appl. Math. **58**, 193–199 (1995)
31. Lin, Y., Stadtherr, M.A.: Deterministic global optimization for parameter estimation of dynamic systems. Ind. Eng. Chem. Res. **45**(4), 8438–8448 (2006)
32. Lin, Y., Stadtherr, M.A.: Deterministic global optimization of nonlinear dynamic systems. AIChE J. **53**(4), 866–875 (2007a)
33. Lin, Y., Stadtherr, M.A.: Validated solutions of initial value problems for parametric ODEs. Appl. Numer. Math. **57**(10), 1145–1162 (2007b)
34. Makino, K., Berz, M.: Remainder differential algebras and their applications. In: Berz, M., Bishof, C., Corliss, G., Griewank, A. (eds.) Computational Differentiation: Techniques, Application, and Tools, pp. 63–75. SIAM, Philadelphia (1996)
35. Makino, K., Berz, M.: Efficient control of the dependency problem based on Taylor model methods. Reliab. Comput. **5**(1), 3–12 (1999)
36. Makino, K., Berz, M.: Taylor models and other validated functional methods. Int. J. Pure Appl. Math. **4**, 379–456 (2003)
37. Mason, J.C., Handscomb, D.C.: Chebyshev Polynomials. CRC Press, Boca Raton (2003)
38. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: part I—convex underestimating problems. Math. Program. **10**, 147–175 (1976)
39. Mitsos, A., Chachuat, B., Barton, P.I.: McCormick-based relaxations of algorithms. SIAM J. Optim. **20**(2), 573–601 (2009)
40. Moore, R.E.: Interval Analysis. Prentice-Hall, Englewood Cliffs (1966)
41. Moore, R.E., Kearfott, R.B., Cloud, M.J.: Introduction to Interval Analysis. SIAM, Philadelphia (2009)
42. Neher, M., Jackson, K.R., Nediakov, N.S.: On Taylor model based integration of ODEs. SIAM J. Numer. Anal. **45**, 236–262 (2007)
43. Neumaier, A.: Taylor forms—use and limits. Reliab. Comput. **9**(1), 43–79 (2002)
44. Neumaier, A.: Complete search in continuous global optimization and constraint satisfaction. Acta Numer. **13**, 271–369 (2004)
45. Park, S.M., Chang, K.Y., Hong, D., Seo, C.: Comments on “on the polynomial multiplication in Chebyshev form”. IEEE Trans. Comput. **63**(12), 3162–3163 (2014)
46. Parrilo, P.A.: Semidefinite programming relaxations for semialgebraic problems. Math. Program. **96**, 293–320 (2003)
47. Paulen, R., Villanueva, M.E., Chachuat, B.: Guaranteed parameter estimation of nonlinear dynamic systems using high-order bounding techniques with domain and CPU-time reduction strategies. IMA J. Math. Control Inf. **33**(3), 563–587 (2015)
48. Rababah, A.: Transformation of Chebyshev–Bernstein polynomial basis. Comput. Methods Appl. Math. **3**(4), 608–622 (1993)
49. Rajyaguru, J., Chachuat, B.: Taylor models in deterministic global optimization for large-scale systems with few degrees of freedom. Comput. Aided Chem. Eng. **32**, 973–978 (2013)
50. Rajyaguru, J., Villanueva, M.E., Houska, B., Chachuat, B.: Continuous-time enclosures for uncertain implicit differential equations. IFAC Pap. OnLine **48**(8), 94–99 (2015)
51. Ratschek, H., Rokne, J.: Computer Methods for the Range of Functions. Series in Mathematics and Its Applications. Ellis Horwood, Chichester (1984)
52. Sahlodin, A.M.: Global Optimization of Dynamic Process Systems Using Complete Search Methods. Ph.D. thesis, McMaster University, Hamilton, ON, Canada (2012)
53. Sahlodin, A.M., Chachuat, B.: Convex/concave relaxations of parametric ODEs using Taylor models. Comput. Chem. Eng. **35**(5), 844–857 (2011)
54. Sherali, H.D., Adams, W.P.: A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems. Kluwer, Boston (1999)
55. Tawarmalani, M., Sahinidis, N.V.: Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications. Series in Nonconvex Optimization and Its Applications. Kluwer, Boston (2002)

56. Townsend, A., Trefethen, L.N.: An extension of Chebfun to two dimensions. *SIAM J. Sci. Comput.* **35**(6), C495–C498 (2013)
57. Trefethen, L.N.: Computing numerically with functions instead of numbers. *Math. Comput. Sci.* **1**, 9–19 (2007)
58. Trefethen, L.N.: *Approximation Theory and Approximation Practice*. SIAM, Philadelphia, PA (2013)
59. Tsoukalas, A., Mitsos, A.: Multivariate McCormick relaxations. *J. Glob. Optim.* **59**(2–3), 633–662 (2014)
60. Villanueva, M.E., Houska, B., Chachuat, B.: Unified framework for the propagation of continuous-time enclosures for parametric nonlinear ODEs. *J. Glob. Optim.* **62**(3), 575–613 (2015a)
61. Villanueva, M.E., Rajyaguru, J., Houska, B., Chachuat, B.: Ellipsoidal arithmetic for multivariate systems. *Comput. Aided Chem. Eng.* **37**, 767–772 (2015b)
62. Wechsung, A., Schaber, S.D., Barton, P.I.: The cluster problem revisited. *J. Glob. Optim.* **58**(3), 429–438 (2014)